

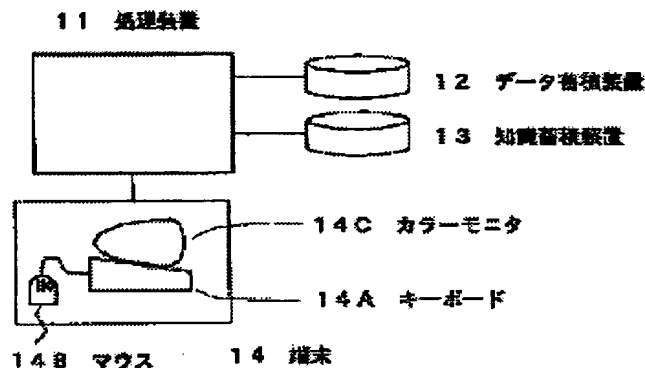
DATA BASE RETRIEVING METHOD

Patent number: JP7234877
Publication date: 1995-09-05
Inventor: TAKAHATA MINORU; others: 03
Applicant: NIPPON TELEGR & TELEPH CORP
Classification:
- International: G06F17/30
- european:
Application number: JP19940025049 19940223
Priority number(s):

Abstract of JP7234877

PURPOSE:To construct a viewpoint corresponding to the intension of a retrieving person and to execute retrieval corresponding to the intension of the retrieving person by using this viewpoint.

CONSTITUTION:At a data base to retrieve data stored in a data storage device 12, a processor 11 dispersedly projects attribute information concerning plural pieces of key data indicated by the retrieving person in a space formed by declaratory knowledge stored in a knowledge storage device 13. This processing uses a parameter given by the retrieving person as the range of the retrieval intension. Next, the processor 11 constructs the viewpoint from the attribute information dispersedly projected in the space formed by the declaratory knowledge. Afterwards, the processor 11 retrieves the data stored in the data storage device 12 by using the constructed viewpoint.



Data supplied from the esp@cenet database - Worldwide

(19)日本国特許庁(JP)

(12)公開特許公報 (A)

(11)特許出願公開番号

特開平 7 - 2 3 4 8 7 7

(43)公開日 平成7年(1995)9月5日

(51)Int. Cl.⁶
G 0 6 F 17/30

識別記号

庁内整理番号

F I

技術表示箇所

9194-5 L

G 0 6 F 15/403 3 2 0 Z

9194-5 L

15/40 3 7 0 B

審査請求 未請求 請求項の数5

O L

(全13頁)

(21)出願番号 特願平6-25049

(22)出願日 平成6年(1994)2月23日

(71)出願人 000004226

日本電信電話株式会社

東京都千代田区内幸町一丁目1番6号

(72)発明者 高畑 実

東京都千代田区内幸町1丁目1番6号 日本
電信電話株式会社内

(72)発明者 坂本 弘章

東京都千代田区内幸町1丁目1番6号 日本
電信電話株式会社内

(72)発明者 上野 圭一

東京都千代田区内幸町1丁目1番6号 日本
電信電話株式会社内

(74)代理人 弁理士 森田 寛 (外1名)

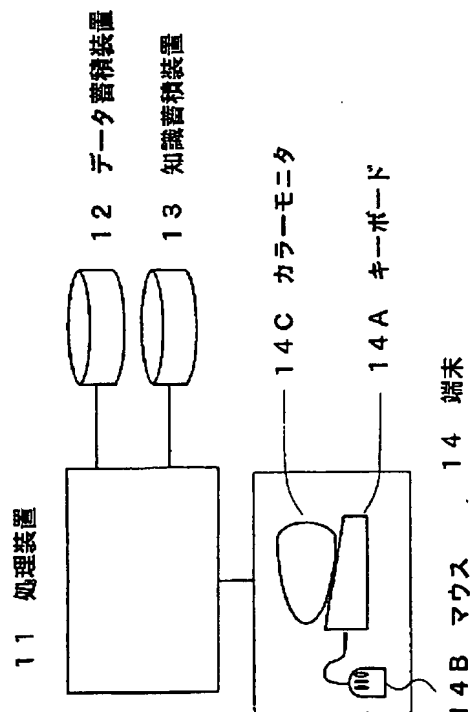
最終頁に続く

(54)【発明の名称】 データベース検索方法

(57)【要約】

【目的】 本発明は、データベース検索方法に関し、検索者の意図に応じて観点を構築し、これを用いて検索者の意図に応じた検索を行うことを目的とする。

【構成】 データ蓄積装置12に蓄積されたデータを検索するデータベースにおいて、処理装置11が、検索者によって指示された複数のキーデータについての属性情報を、知識蓄積装置13に格納された宣言的知識によって形成される空間へ拡散投影する。この処理は、検索者が検索意図の範囲として与えるパラメータを用いる。次に、処理装置11が、宣言的知識によって形成される空間へ拡散投影された属性情報から、観点を構築する。次に、処理装置11が、構築された観点をを用いて、データ蓄積装置12に蓄積されたデータを検索する。



【特許請求の範囲】

【請求項 1】 データを蓄積し、管理し、蓄積されたデータを検索するデータベース検索システムにおいて、検索者によって指示された複数のキーデータについての属性情報を、宣言的知識によって形成される空間へ拡散投影する第 1 の処理と、宣言的知識によって形成される空間へ拡散投影された属性情報から観点を構築する第 2 の処理とを有することを特徴とするデータベース検索方法。

【請求項 2】 前記第 1 の処理は検索者が検索意図の範囲として与えるパラメータを用いることを特徴とする請求項 1 に記載のデータベース検索方法。

【請求項 3】 前記第 1 及び第 2 の処理によって構築された観点をを用いて、前記空間において検索対象のデータの属性情報が占める位置における前記観点の値を求めることにより、当該検索対象のデータの検索評価値を求めてデータベースを検索することを特徴とする請求項 1 又は請求項 2 に記載のデータベース検索方法。

【請求項 4】 前記データベース検索システムは画像データを蓄積し、管理し、蓄積された画像データを検索するものであって、前記第 1 の処理において、検索者によって指示された複数のキーデータである候補画像データについての属性情報である特徴量属性を、宣言的知識である特徴量属性の関係によって形成される空間へ拡散投影し、前記第 2 の処理において、特徴量属性の関係によって形成される空間へ拡散投影された特徴量属性から、画像データの検索の観点を構築することを特徴とする請求項 1 に記載のデータベース検索方法。

【請求項 5】 前記データベース検索システムはキーワードデータを蓄積し、管理し、蓄積されたキーワードデータを検索するものであって、前記第 1 の処理において、検索者によって指示された複数のキーデータであるキーワードデータについての属性情報である意味属性を、宣言的知識である意味属性の関係によって形成される空間へ拡散投影し、前記第 2 の処理において、前記意味属性の関係によって形成される空間へ拡散投影された意味属性から、キーワードデータの検索の観点を構築することを特徴とする請求項 1 に記載のデータベース検索方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、データベース検索方法に関し、特に、検索者によって指示された複数のキーデータから検索者がどの特徴を重視しているのかという観点を構築し、構築された観点を元にデータの検索を行うデータベース検索方法に関する。

【0002】

【従来の技術】一般にデータの類似性は、検索者がどのような特徴を重視しているのかという点（観点）に大きく依存する。このため、類似するデータを多数蓄積、管

理しているデータベースを検索するにあたっては、検索者の観点を考慮することが望ましい。

【0003】従来の検索者の観点によるデータベースの検索方法は、例えば、沢田、大川、馬場口、手塚らによる「観点を考慮した連想機構の一モデル化」、情報処理学会研究報告 情報学基礎、28-2, pp9-16, (Nov. 1992) に示されている。これによれば、検索者によって提示された複数のキーデータがどのような特徴に関して類似しているのか（検索者の検索意図）を、システム側がデータベース中で評価することによって、どこに重みを付与して検索すれば良いのか（観点）という形で算出する。

【0004】

【発明が解決しようとする課題】しかしながら、前述した従来のデータベース検索方法によれば、検索者によって提示された複数のキーデータがどのような特徴に関して類似しているのか（検索者の検索意図）の評価は、システムが、当該複数のキーデータの各々の属性情報について、データベースの全体の中に占める割合という相対的な尺度を基に計算することにより求めている。例えば、ある属性情報に着目して、同一の属性のものがデータベース中に少なければそのデータは特殊性が高く、多ければそのデータは一般性が高い。そこで、提示されたキーデータの属性情報の各々について、データベース中の同一属性のものの割合を求め、次に、特殊性が高い属性情報の重みを大きくし一般性が高い属性情報の重みを小さくして、データベースを検索する。

【0005】このため、当該複数のキーデータ毎の属性テーブル（属性情報を格納するテーブル）が、予め同一の構成になっている必要がある。即ち、データベース中のデータの各々の属性テーブルが、同一の属性情報を格納している必要がある。もし、格納していなければ、データベースの全体の中に占める割合を求めることができない。しかし、多種多様なデータを扱うデータベースにおいては、格納しているデータの各々について、一貫した属性テーブルを予め用意することは困難であるという問題がある。

【0006】また、前述した従来のデータベース検索方法によれば、検索者の検索意図を、データベースの全体の中に占める割合という相対的な尺度を基に計算して評価している。しかし、検索者は、そのキーデータの属性情報の各々と同一属性のものがデータベース中で占める割合を、キーデータを提示する時点で知る手段を持たない。従って、検索者の検索意図は、本来キーデータがどの程度の手掛かりを持つのかを示すはずであるのに、これがデータベースの内容によって決定されてしまい、検索者の検索意図に沿わなくなるという不都合を生じる。

【0007】このため、検索者は、自己の検索意図及びその範囲をシステム側に明示的に与える手段を持たなかった。従って、また、これにより構築される観点を自己

の検索意図に明確に沿ったものとする手段も持たなかった。このため、検索結果を特定しすぎたり、逆に、検索結果を絞りきれないと言う問題があった。

【0008】本発明の目的は、検索者の意図に応じて観点を構築し、これを用いて検索者の意図に応じた検索を行うことが可能なデータベース検索方法を提供することにある。

【0009】また、本発明の他の目的は、データ毎に属性テーブルに一貫性が保証されていないような多種多様なデータを扱うデータベースにおいて、検索者の観点によって検索結果として期待するデータが変化しても対応可能なデータベース検索方法を提供することにある。

【0010】更に、本発明の他の目的は、検索者の検索意図の範囲を明示的に与え、検索の精度と作業効率を向上することが可能なデータベース検索方法を提供することにある。

【0011】

【課題を解決するための手段】本発明は、検索者によって指示されたキーデータに基づいて、各々の属性情報を用意し、各々の属性情報に対応する種類の知識を用いて観点を構築し、構築された観点を基にデータ蓄積装置に蓄積されたデータの検索を行うデータベース検索方法であって、知識を用いて属性情報の異なるデータ群から観点を構築することを主要な特徴とする。

【0012】即ち、知識を用いて属性情報の異なるデータ群から観点を構築する過程は、検索者によって指示された複数のキーデータについての属性情報を、宣言的知識によって形成される空間へ拡散投影する第1の処理と、宣言的知識によって形成される空間へ拡散投影された属性情報から観点を構築する第2の処理とを有する。また、前記第1の処理は、検索者が検索意図の範囲として与えるパラメータを用いる。

【0013】

【作用】本発明によれば、検索者によって提示された複数のキーデータについての属性情報を用意した後、これを用いて、宣言的知識により形成される空間上へ観点が構築される。

【0014】即ち、観点が、検索者によって提示された複数のキーデータについての各々の属性情報がデータベースの全体の中に占める割合という相対的な尺度を基に求められることは無い。これにより、検索者は、自己の検索意図、従って観点を、データベースの内容とは無関係に構築することができる。従って、検索者の意図に応じて観点を構築し、所望の検索結果を得ることができる。

【0015】また、観点が、属性情報を直接計算に用いることによって求められることは無い。これにより、データ毎に属性テーブルに一貫性が保証されていないような多種多様なデータを扱うデータベースであっても、検索者の観点を求めることができる。従って、検索者の観

点によって検索結果として期待するデータが変化しても対応可能であり、所望の検索結果を得ることができる。

【0016】また、本発明によれば、宣言的知識によって形成される空間へ観点を構築する過程では、検索者の検索意図の範囲をパラメータとして与えることができる。これにより、検索者は、自己の検索意図をより明確に反映させた観点を構築することができるので、検索の精度と作業効率を向上することができる。

【0017】

【実施例】以下、図面を用いて、本発明の実施例を詳細に説明する。図1は実施例構成図であり、本発明によるデータベース検索方法を実施するデータベース検索システムの構成を示すブロック図である。

【0018】このデータベース検索システムは、処理装置11、データ蓄積装置12、知識蓄積装置13、端末14等により構成される。これらの各システム構成要素はLANやWAN等のネットワークを介して接続されていても良い。

【0019】処理装置11は、データベースの検索処理等を行うものであり、CPU（中央処理装置）とメモリとからなり、メモリ上に当該データベースの管理、検索を行うデータベース管理処理プログラムを有する。データ蓄積装置12は、検索対象である画像やテキスト等のデータと、必要に応じて各データについての属性情報とを格納する。知識蓄積装置13は宣言的知識である属性情報の関係情報等を格納する。端末14は、検索者が検索条件（キーデータ、パラメータ）等を入力するためのキーボード14A、マウス14B、検索結果等を表示するためのカラーモニタ14C等のユーザインタフェースを司る手段を有する。

【0020】図2は、本発明の主要な特徴である属性情報の異なるデータ群から、知識を用いて観点を構築する原理を表した図である。図2において、21は検索者によって指示されたキーデータ（検索キー）、22はキーデータ21の属性情報、23は宣言的知識によって形成される空間、24は属性情報22の拡散投影によって生成されたフィールド、25は属性情報22をどの程度拡散するかを定めるためのパラメータであって検索者が検索意図の範囲として与えるもの、26は多数のフィールド24によって形成される観点、27は属性情報22の空間23における位置を表す。

【0021】なお、ここで、拡散投影とは、ある点の近傍において任意の2点間の距離が定義できる空間に、ある情報が投影される場合、その距離に従って拡散したフィールドを生成する演算をいう。

【0022】また、宣言的知識は、属性情報の関係等を定め、この関係等を宣言的に記述しており、例えば宣言的に記述したルールからなる。なお、宣言的知識については、例えば、「認識と学習」（安西著、岩波講座、ソフトウェア科学16、岩波書店、1989年発行）に示

されている。

【0023】図2において、検索者によって指示されたキーデータ21はD1、D2の2個である。キーデータD1の属性情報22は、Pa1、Pa2、Pb1、Pc1、Pd1であり、キーデータD2の属性情報22は、Pa3、Pc2、Pe1、Pf1である。従って、各キーデータ21の属性情報22は互いに異なる。しかし、本発明によれば、属性情報22が直接（検索評価のための）計算に用いられる事がなく、宣言的知識により一旦空間23上に投影されて用いられるので、属性情報22が異なっている問題ない。

【0024】宣言的知識によって、属性情報22を投影するための空間23が規定される。この空間23上に、宣言的知識を用いることによって、各キーデータ21の属性情報22の拡散投影であるフィールド24が生成される。各フィールド24の範囲、即ち、検索者の検索意図の範囲は、属性情報22の空間23における位置27を中心としてパラメータ25によって規定される（円形に限られない）。そして、観点26は多数のフィールド24によって形成される。

【0025】以上から判るように、観点26がどのように構築されるかは、キーデータ21がどれであるか、及び、パラメータ25の値がどの程度であるかによって決まる。本発明によれば、キーデータ21及びパラメータ25の値は検索者が指示できるから、検索者の検索意図に沿った観点26を構築することができる。

【0026】図3は、構築された観点をを用いてデータを検索する原理を表した図である。図3において、31はデータ蓄積装置12から読み出されたデータ、32はデータ31の属性情報、33は属性情報32の空間23における位置、34は位置33における観点26の値を表す。ここで、属性情報32は、データを検索するためのキーデータ等の索引情報の他に、データから抽出された特徴量であってもよい。

【0027】図3において、データ蓄積装置12から検索対象であるデータ31としてデータD3が読み出される。データD3の属性情報32は、Pa4、Pa5、Pc3、Pc4、Pg1である。データD3の属性情報32は、前述の各キーデータD1、D2の属性情報22と異なる。

【0028】一方、図2に示したように、宣言的知識によって空間23が規定され、この空間23上に、宣言的知識を用いることによって、検索者が指示したキーデータ21及びパラメータ25の値に従った観点26、即ち、検索者の検索意図に沿った観点26が構築されている。

【0029】データD3の属性情報Pa4が空間23上で占める位置33は、空間23を規定する宣言的知識と属性情報Pa4とから定まる。この位置33を用いて、観点26の位置33における値34を求める。同様の処

理を各属性情報Pa5、Pc3、Pc4、Pg1について行う。最後に、各属性情報Pa4、Pa5、Pc3、Pc4、Pg1から求まる値34の和を求め、これを検索評価値とする。検索評価値が大きければ、観点26の値34が大きいのであるから、検索者の検索意図に沿ったデータと言える。検索評価値が小さければ、観点26の値34が小さいのであるから、検索者の検索意図に沿わないデータと言える。

【0030】データベース（データ蓄積装置12）中の各データについて検索評価値を求め、例えばその大きい順にソートすることにより、検索者の検索意図に沿ったデータを検索により求めることができる。

【0031】（例1）本発明の一例として、検索対象データが画像である場合について説明する。この場合、データ蓄積装置12は画像情報を格納する画像情報蓄積装置である。ここで、画像情報は、画像と、これについての属性情報である特徴量属性（例えば、色成分に分解したヒストグラム）とからなる。特徴量属性は、例えば処理装置11によって画像から解析されるものである。また、知識蓄積装置13は特徴量属性の関係（例えば、色の近さを宣言的に記述したルール）と、画像から特徴量属性を求めるメソッドとを格納する。このメソッドを処理装置11によって実行することにより、画像から特徴量属性が求まる。

【0032】図4はこの例におけるカラーモニタ14Cの表示画面の一例である。図4において、41は候補画像42を表示するためのウィンドウ、42はキーデータの候補である候補画像、43は候補画像42を指示するためのポインタ、44はルートウィンドウである。なお、ここで、候補画像42は、データ蓄積装置12に格納されていない画像、例えばデータ蓄積装置12に格納された画像のいくつかを合成した画像や、カメラ入力 of 画像であっても良い。

【0033】まず、本実施例のデータベース検索方法における観点構築時の処理の流れについて、図1、図2、図4、図5に従って説明する。図5において、観点構築処理が開始され、処理装置11が図4のようにいくつかの候補画像（キーデータの候補）を端末14に表示する。なお、パラメータ25は、例えば予め与えられる。

【0034】検索者が、端末14からポインタ43を用いて、表示された複数の候補画像42の中から自己の希望する画像に近い1又は2以上の候補画像42（キーデータ）を指示する（S1）。

【0035】これに応じて、処理装置11が、指示された候補画像42についての特徴量属性（属性情報）がデータ蓄積装置12に格納されているか否かを確認する（S2）。

【0036】特徴量属性が格納されていない場合、処理装置11が、知識蓄積装置13に格納されたメソッドによって当該指示された候補画像42を解析して、その特

微量属性を求める (S3)。

【0037】一方、特徴量属性が格納されている場合、処理装置11が、データ蓄積装置12からこれを読み出す (S4)。なお、ここで、検索者が、指示した候補画像42の特徴量属性を端末14から直接入力しても良い (S5)。

【0038】次に、処理装置11が、知識蓄積装置13に格納された特徴量属性の関係 (宣言的知識) に基づいて形成される空間23へ、特徴量属性を拡散投影し、フィールド24を得る (S6)。この拡散投影は、検索者が検索意図を示す範囲として与えるパラメータ25によって決められる範囲に行われる。この例の場合の空間23は、図6に示すような、例えば色成分及び頻度を軸*i*及び*h*として持つ空間である。

【0039】次に、処理装置11が、それまでに得たフィールド24を積算して、観点26を求める (S7)。次に、処理装置11が、S1で指定された候補画像42 (キーデータ) が残っているか否かを調べる (S8)。

【0040】残っている場合、以上の処理を各候補画像42について繰り返し、残っていない場合、処理を終了する。図6は、以上の処理によって構築された観点の例を示す。

【0041】図6において、Cは (特徴量属性の) 観点、*i* は色成分を示す軸、*h* は頻度を示す軸であり、R、G、Bは色成分である。この例の場合の空間23は、図6に示すように、例えば色成分軸*i*と頻度軸*h*とを持つ空間であり、色の近さを宣言的に記述したルールである宣言的知識により定まる。また、この例の場合の特徴量属性は色成分に分解したヒストグラムからなるので、色成分と頻度とをそれぞれ軸に持つ2次元空間における特徴量属性の位置が定まる。そこで、宣言的知識を用いて、特徴量属性を空間23へ拡散投影することにより、フィールドが求まる。そして、多数のフィールドの積算により、観点Cが求まる。図6において、色成分軸*i*と頻度軸*h*とを持つ空間上に形成された山状の図形の集合が観点Cであり、その高さ (図中、Z軸方向の高さ) が観点Cの値である。

【0042】次に、本実施例により構築した観点を用いたデータベース検索時の処理の流れについて、図1、図3、図4、図7に従って説明する。図7において、検索処理が開始され、処理装置11が、画像とその特徴量属性 (属性情報) とからなる画像情報を蓄積するデータ蓄積装置12から、1個の画像 (検索対象データ) を読み出す (S9)。

【0043】さらに、処理装置11が、データ蓄積装置12から、S9で読み出した画像に対応する1個の特徴量属性を読み出す (S10)。次に、処理装置11が、知識蓄積装置13に格納された特徴量属性の関係 (宣言的知識) に基づいて形成される空間23において、S10で読み出した特徴量属性の位置33を定め、位置33

から観点26の値34 (観点26が位置33において取る値34) を求める (S11)。

【0044】次に、処理装置11が、S11で得た値34をそれまでに得た観点26の値34 (検索評価値) に加算して、新しい検索評価値とする。 (S12)。検索評価値は、画像 (検索対象データ) ごとに求められる。

【0045】次に、処理装置11が、S9で読み出した画像に対応する特徴量属性が残っているか否かを調べる (S13)。残っている場合、S10以下の処理を各特徴量属性について繰り返す。

【0046】残っていない場合、処理装置11が、検索範囲にある画像 (検索対象データ) がデータ蓄積装置12に残っているか否かを調べる (S14)。残っている場合、S9以下の処理を各特徴量属性について繰り返す。なお、検索範囲は、検索者が入力しても、また、データベース又はデータ蓄積装置12内の所定のデータについて行うように予め定めても良い。

【0047】残っていない場合、処理装置11が、検索評価値を調べて、これが大きい順に画像 (検索対象データ) をソートして (S14)、処理を終了する。ここで、図6において、観点Cを用いた以上の処理によって、画像の検索を行う場合の例を示す。

【0048】前述のように、この例の場合の特徴量属性は色成分に分解したヒストグラムからなるから、色成分と頻度とをそれぞれ軸に持つ2次元空間において、特徴量属性の位置が定まる。そこで、観点Cの当該位置における値 (図中、Z軸方向の高さ) を求め、検索評価値を求める。これにより、検索意図に沿った検索結果を得ることができる。

【0049】以上の説明から判るように、本発明によれば、データベースの内容とは無関係に観点を構築でき、検索キー、又は、データベース中の画像ごとに特徴量属性の一貫性が保証されていない場合においても、検索者の観点を計算できる。従って、検索者の観点によって検索結果として期待される画像が変化しても、これに対応した結果を得る検索が可能である。

【0050】また、本発明によれば、パラメータ25によって、検索者の検索意図の範囲を明示的に与えることができるので、検索の精度と作業効率を向上することができる。

【0051】 (例2) 次に、本発明の他の一例として、検索対象データが文字である場合について説明する。この場合、データ蓄積装置12は文字情報を格納する文字情報蓄積装置である。ここで、文字情報は、キーワード (文字) と、これについての属性情報である意味属性 (例えば、「りんご」であれば、「実」、「果物」、「果樹」等、単語の見方を整理したもの) とからなり、キーワードとの意味的な距離に従った値を持つ。意味属性は、例えば処理装置11によってキーワードから解析されるものである。また、知識蓄積装置13は意味属性

10

20

30

40

50

の関係と、シソーラスと、意味属性辞書とを格納する。意味属性の関係は、宣言的知識であって、例えば、意味属性をノードとして宣言的な上位概念、下位概念等のリンクで結合された、図9に示すような意味ネットワーク81からなる。シソーラスは、キーワードの同義語、類義語と、キーワードから意味属性を求めるための意味的な距離に従った値とを有する。意味属性辞書はキーワードの意味属性を格納する。

【0052】図9は、この例における観点の構築前の空間23と構築後の空間23を示す。特に、図9の上側は観点の構築前の空間23を示し、図9の下側は観点の構築後の空間23を示す。

【0053】図9において、81は意味属性をノードとして宣言的な上位概念、下位概念等のリンクで結合された意味ネットワーク、82は初期ノード（図中、黒丸で示す）、83は初期ノード82によって活性伝播した後の活性パターン（図中、太線で示す）である。

【0054】ここでは、図9の上側に示した観点の構築前の空間23について説明する。この例の場合の空間23は、宣言的知識である意味属性の関係によって形成される意味ネットワーク81からなる。意味ネットワーク81はノードとリンクとからなる。ノードは意味属性であり、図中白丸及び黒丸（初期ノード82）で示される。リンクは意味属性であるノードを宣言的な上位概念、下位概念等を結合するものであり、図中ノードを結ぶ点線で示される。リンクの距離は、シソーラスの格納する意味的な距離に従った値から定まる。

【0055】まず、本実施例のデータベース検索方法における観点構築時の処理の流れについて、図1、図2、図4、図8に従って説明する。図8において、観点構築処理が開始され、処理装置11がいくつかの候補ワード（キーワードの候補）を端末14に表示する。なお、パラメータ25は、例えば予め与えられる。

【0056】検索者が、端末14から前述のポインタ43と同様の手段を用いて、表示された複数の候補ワードの中から自己の希望するデータに近い1又は2以上のキーワードを指示する（S16）。なお、端末14からキーワードそれ自体を直接入力しても良い。ここで、検索者が指示したキーワードの意味属性が、図9に示す初期ノード82であり、拡散投影（実際は活性伝播）の初期値とされる。

【0057】これに応じて、処理装置11が、指示されたキーワードの意味属性（属性情報）がデータ蓄積装置12に格納されているか否かを確認する（S17）。意味属性が格納されていない場合、処理装置11が、知識蓄積装置13に格納されたシソーラスから、入力されたキーワードの同義語、類義語を読み出す（S18）。ここで、読み出された同義語、類義語は、キーワードとの意味的な距離に従った値を有する。この同義語、類義語は、検索者の入力したキーワードに代わるものとして用

いられる。

【0058】次に、処理装置11が、知識蓄積装置13に格納された意味属性辞書から、S18で読み出された同義語、類義語の意味属性を読み出す（S19）。ここで、読み出された意味属性は、キーワードとの意味的な距離に従った値を有する。この同義語、類義語の意味属性は、検索者の入力したキーワードの意味属性に代わるものとして、図9に示す初期ノード82とされ、拡散投影（実際は活性伝播）の初期値として用いられる。

【0059】一方、意味属性が格納されている場合、処理装置11が、データ蓄積装置12からこれを読み出す（S20）。なお、ここで、検索者が、指示したキーワードの意味属性を端末14から直接入力しても良い（S21）。

【0060】次に、処理装置11が、知識蓄積装置13に格納された意味属性の関係（宣言的知識）に基づいて形成される空間23である意味ネットワーク81へ、意味属性を拡散投影し、フィールド24を得る（S22）。

【0061】ここで、本例では、拡散投影の手段として、その下位概念である活性伝播を用いる。即ち、本例の図9においては、ある情報が空間23（意味ネットワーク81）上に投影された場合に、フィールド24を生成する演算は、2次元方向ではなく、意味ネットワーク81のリンクの方向、即ち1次元方向に行われる。活性伝播は、検索者が検索意図を示す範囲として与えるパラメータ25によって決まる伝播係数で定まる範囲に行われ、これによりフィールド24がリンクの方向に沿って広がって定まる。また、活性伝播は、検索者が指示したキーワードの意味属性をその初期値（初期ノード82）として行われる。

【0062】次に、処理装置11が、それまでに得たフィールド24を積算して、観点26を求める（S23）。次に、処理装置11が、S16で指定されたキーワードが残っているか否かを調べる（S24）。

【0063】残っている場合、以上の処理を各キーワードについて繰り返し、残っていない場合、処理を終了する。ここで、以上の処理によって空間23である意味ネットワーク81に構築された観点であって、図9の下側に示した観点について説明する。

【0064】この例の場合、観点は、初期ノード82（図中、黒丸で示す）によって活性伝播した後の活性パターン83（図中、太線で示す）である。また、この例の場合の空間23は、前述したように、宣言的知識である意味属性の関係によって形成される意味ネットワーク81からなり、意味属性をノードとし、上位概念、下位概念のリンクで結合されている。また、この例の場合の意味属性は、意味ネットワーク81からなる空間23において、その位置が定まる。そこで、宣言的知識を用いて、意味属性を空間23へ活性伝播することにより、フ

ィールド 24 が求まる。リンクの距離（長さ）はセンサーの格納する意味的な距離に従った値から定まっており、フィールド 24 の範囲（活性伝播の距離）は、パラメータ 25 によって定まる伝播係数で決定される。そして、多数のフィールド 24 の積算により、観点である活性パターン 83 が求まる。図 9 において、意味ネットワーク 81 のリンクに沿って形成された実線が観点である活性パターン 83 であり、活性パターン 83 の線の太さ（幅）が観点の値である。

【0065】次に、本実施例により構築した観点を用いたデータベース検索時の処理の流れについて、図 1、図 3、図 4、図 10 に従って説明する。図 10 において、検索処理が開始され、処理装置 11 が、キーワードとその意味属性（属性情報）とからなる文字情報を蓄積するデータ蓄積装置 12 から、1 個のキーワード（検索対象データ）を読み出す（S25）。

【0066】さらに、処理装置 11 が、データ蓄積装置 12 から、S25 で読み出したキーワードに対応する 1 個の意味属性を読み出す（S26）。次に、処理装置 11 が、知識蓄積装置 13 に格納された意味属性の関係（宣言的知識）に基づいて形成される空間 23 である意味ネットワーク 81（又は、活性パターン 83）において、S26 で読み出した意味属性であるノードの位置を求め、このノードに接続された観点 26 である活性パターン 83 の値 34 を求める（S27）。

【0067】次に、処理装置 11 が、S27 で得た値 34 をそれまでに得た観点 26 の値 34（検索評価値）に加算して、新しい検索評価値とする。（S28）。検索評価値は、キーワード（検索対象データ）ごとに求められる。

【0068】次に、処理装置 11 が、S25 で読み出したキーワードに対応する意味属性が残っているか否かを調べる（S29）。残っている場合、S26 以下の処理を各意味属性について繰り返す。

【0069】残っていない場合、処理装置 11 が、S28 で求めた検索評価値を、検索範囲にあるキーワード（検索対象データ）の数で除算して、正規化する（S30）。なお、検索範囲は、検索者が入力しても、また、データベース又はデータ蓄積装置 12 内の所定のデータについて行うように予め定めても良い。

【0070】次に、処理装置 11 が、検索範囲にあるキーワード（検索対象データ）がデータ蓄積装置 12 に残っているか否かを調べる（S31）。残っている場合、S25 以下の処理を各特徴量属性について繰り返す。

【0071】残っていない場合、処理装置 11 が、S30 で正規化された検索評価値を調べて、これが大きい順にキーワード（検索対象データ）をソートして（S32）、処理を終了する。

【0072】ここで、図 9 において、以上の処理によって観点である活性パターン 83 を用いて検索を行う場合

の例を示す。前述のように、この例の場合、ある意味属性が指定されると、その意味ネットワーク 81 及び活性パターン 83 において、意味属性即ちノードの位置が定まる。そこで、このノードに接続された観点である活性パターン 83 の値（接続された各々の線分の太さの和）を求め、検索評価値を求める。これにより、検索意図に沿った検索結果を得ることができる。

【0073】以上の説明から判るように、本発明によれば、データベースの内容とは無関係に観点を構築でき、検索キー、又は、データベース中のデータ毎に意味属性の一貫性が保証されていない場合においても、検索者の観点を計算できる。従って、検索者の観点によって検索結果として期待されるキーワードが変化しても、これに対応した結果を得る検索が可能である。

【0074】また、本発明によれば、パラメータ 25 によって、検索者の検索意図の範囲を明示的に与えることができるので、検索の精度と作業効率を向上することができる。

【0075】以上のように、本発明によれば、データ毎に属性テーブルの一貫性が保証されていない多種多様なデータを扱うデータベースにおいても、検索者の重視している特徴、即ち観点を算出することができるので、検索者の検索意図を反映した検索結果を得ることが可能となる。

【0076】ここで、観点の構築の様子を図 11 に従って説明する。データベースに格納されたデータである画像の中から、キーデータ（検索キー）として画像 91 及び画像 92 を提示した場合を考える。この場合、双方に長方形 94 と三角形 95 とが存在するので、観点として、長方形と三角形とが得られる。この時に得られる観点は、宣言的知識によって得られる距離と、検索者の検索意図が示すパラメータに従った値を持つ。また、キーデータとして画像 92 及び画像 93 を提示した場合、観点は星状図形 96 と凹状図形となる。なお、ここでは形状を用いて説明したが、色等の他の物理的な特徴や意味であっても良い。

【0077】また、検索者の検索意図の範囲をパラメータによって明示的に与えることができるので、検索の精度と作業効率を向上することが可能となる。

【0078】

【発明の効果】以上説明したように、本発明によれば、データ毎に属性テーブルの一貫性が保証されていない多種多様なデータを扱うデータベースにおいても、検索者の重視している特徴、即ち観点を算出することができるので、検索者の検索意図を反映した検索結果を得ることが可能となる。

【0079】また、本発明によれば、データベース検索処理において、データベースの内容とは無関係に観点を構築でき、検索キー、又は、データベース中のデータ毎に意味属性の一貫性が保証されていない場合において

13

も、検索者の観点を計算で求めることができる。従って、検索者の観点によって検索結果として期待されるデータが変化しても、これに対応した結果を得る検索が可能である。

【0080】更に、本発明によれば、パラメータによって、検索者の検索意図の範囲を明示的に与えることができるので、検索の精度と作業効率を向上することができる。

【図面の簡単な説明】

【図1】本発明のデータベース検索方法を実施するデータベース検索システムを示すブロック図である。

【図2】属性の異なるデータ群から知識を用いて観点を構築する原理を表す図である。

【図3】構築された観点をを用いてデータを検索する原理を表す図である。

【図4】一実施例におけるカラーモニタの表示画面の例を示す図である。

【図5】本発明の一実施例における処理の流れを示すフローチャートである。

14

【図6】本発明の一実施例において構築した観点を示す図である。

【図7】本発明の一実施例における処理の流れを示すフローチャートである。

【図8】本発明の一実施例における処理の流れを示すフローチャートである。

【図9】本発明の一実施例における意味ネットワーク上の活性伝播を示す図である。

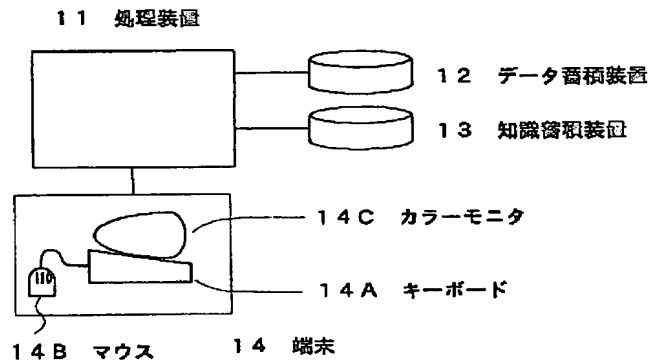
【図10】本発明の一実施例における処理の流れを示すフローチャートである。

【図11】本発明の効果を説明する図である。

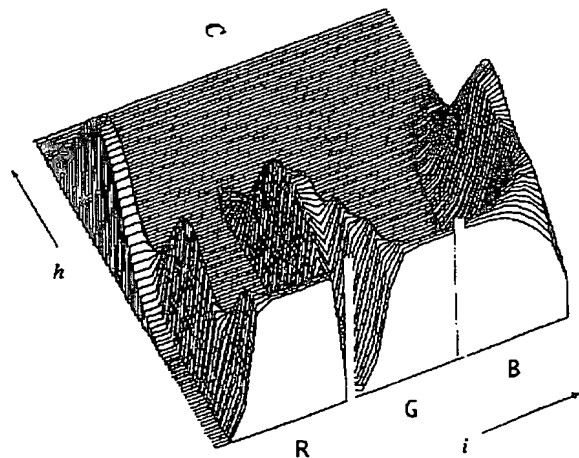
【符号の説明】

- 11 処理装置
- 12 データ蓄積装置
- 13 知識蓄積装置
- 14 端末
- 14 A キーボード
- 14 B マウス
- 14 C カラーモニタ

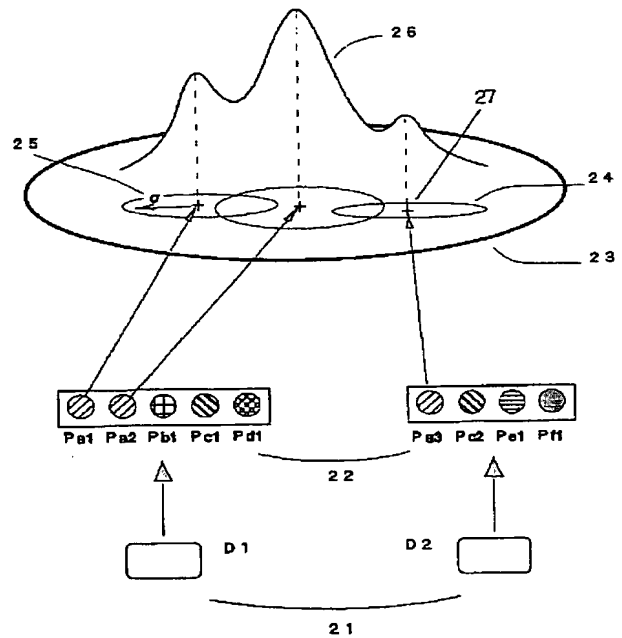
【図1】



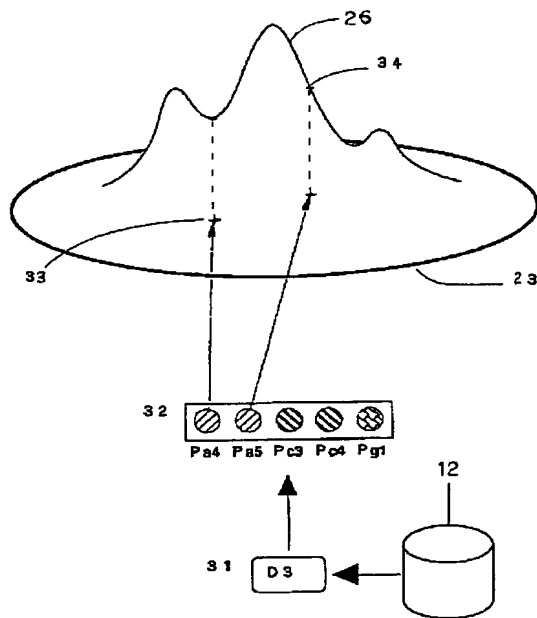
【図6】



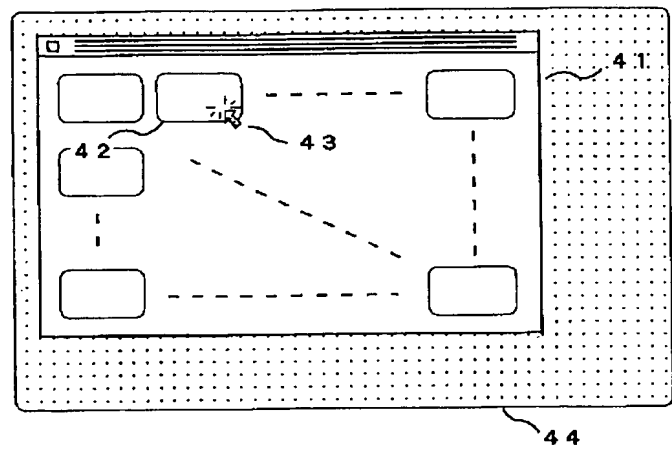
【図2】



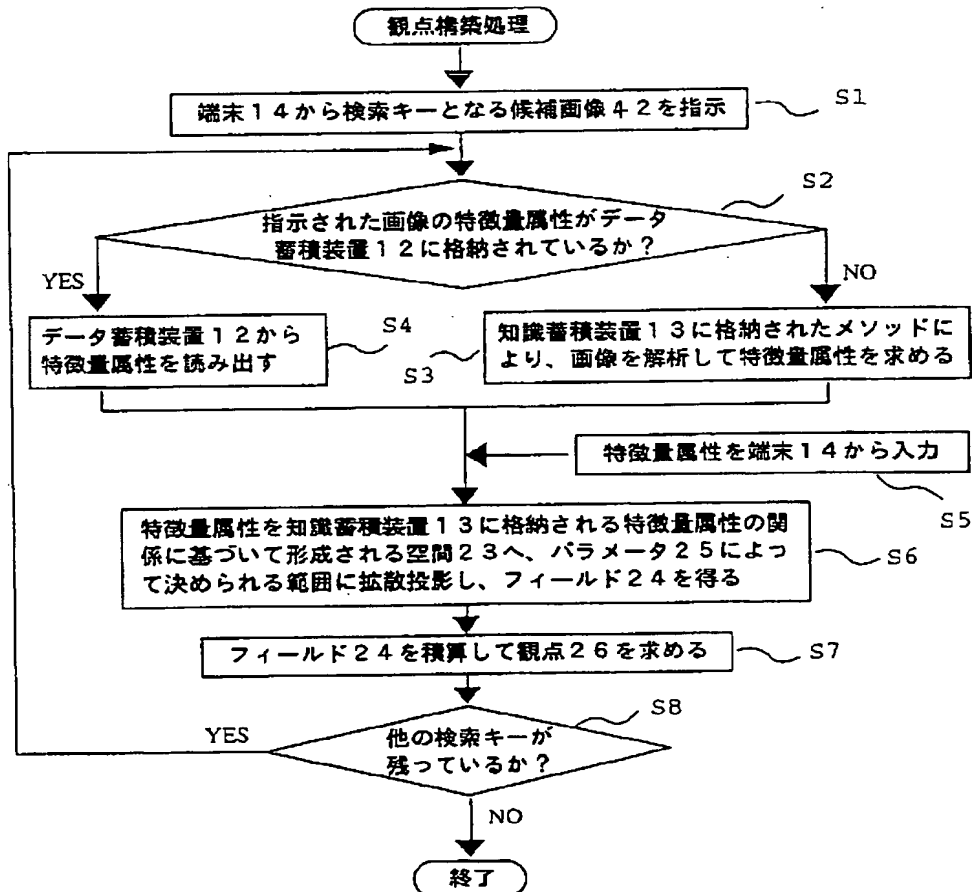
【図3】



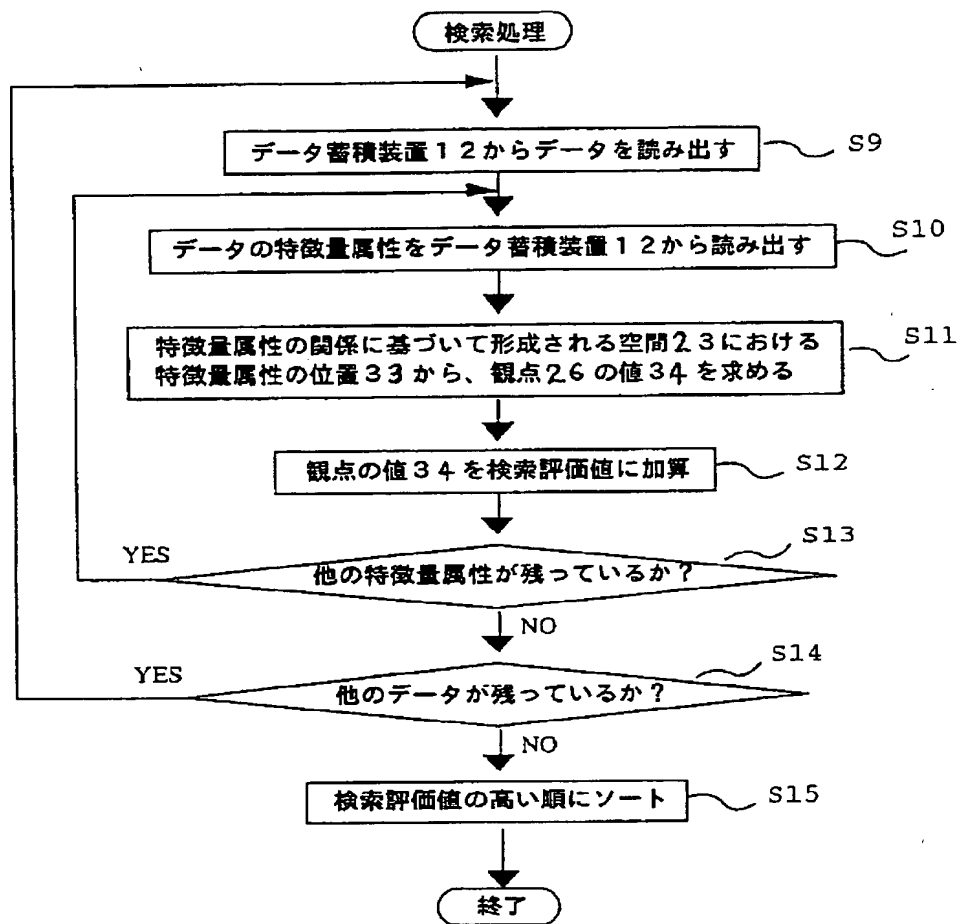
【図4】



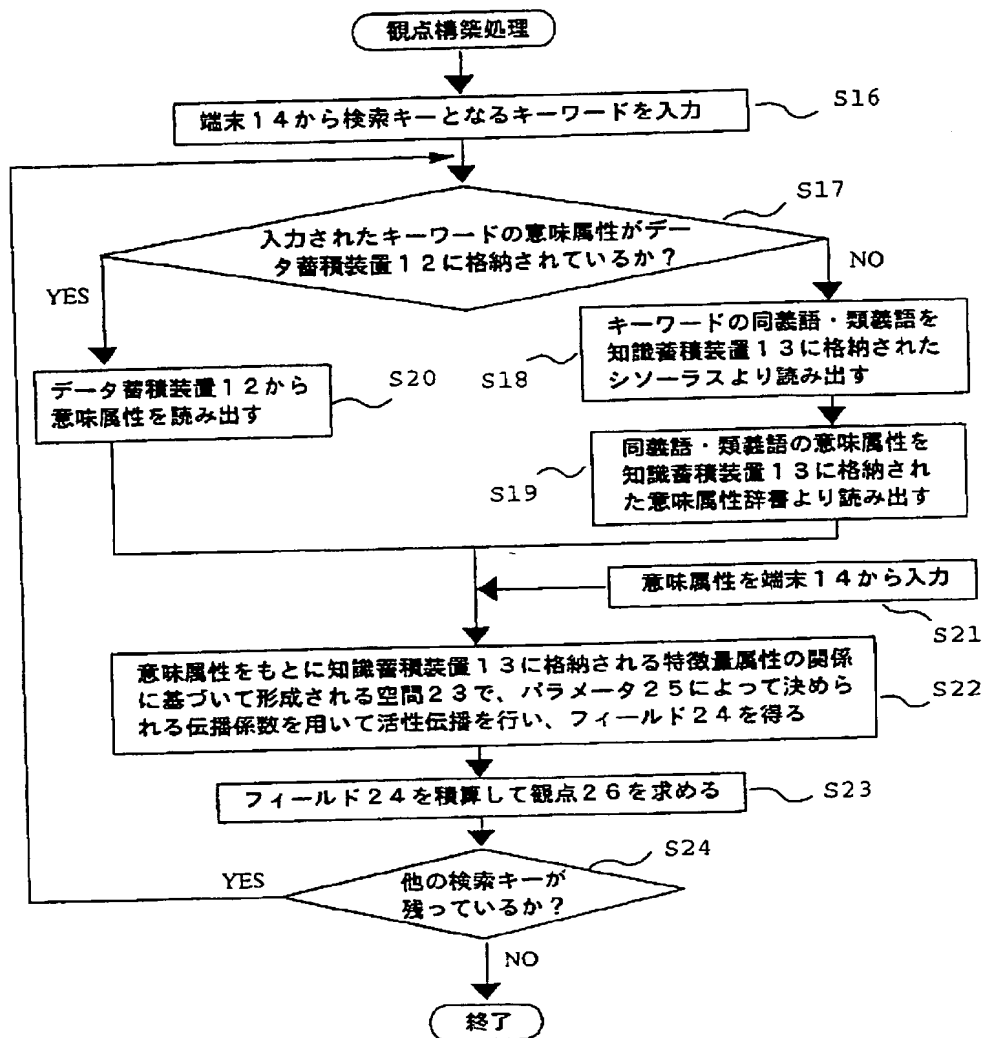
【図5】



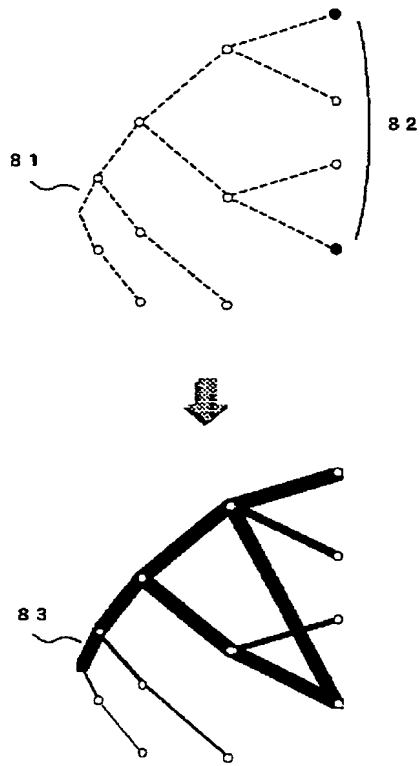
【図 7】



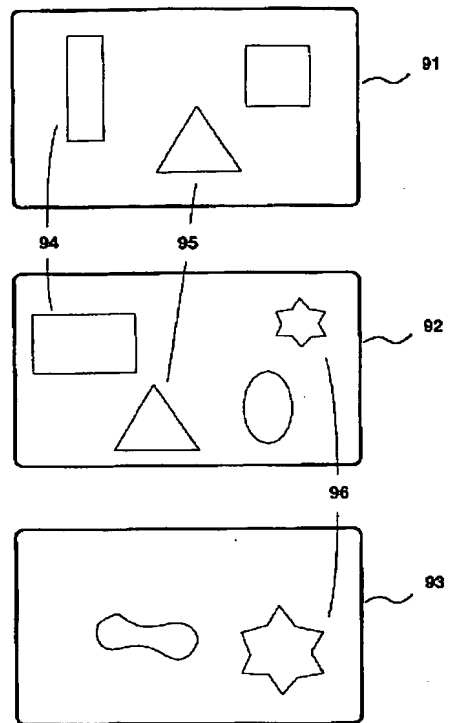
【図 8】



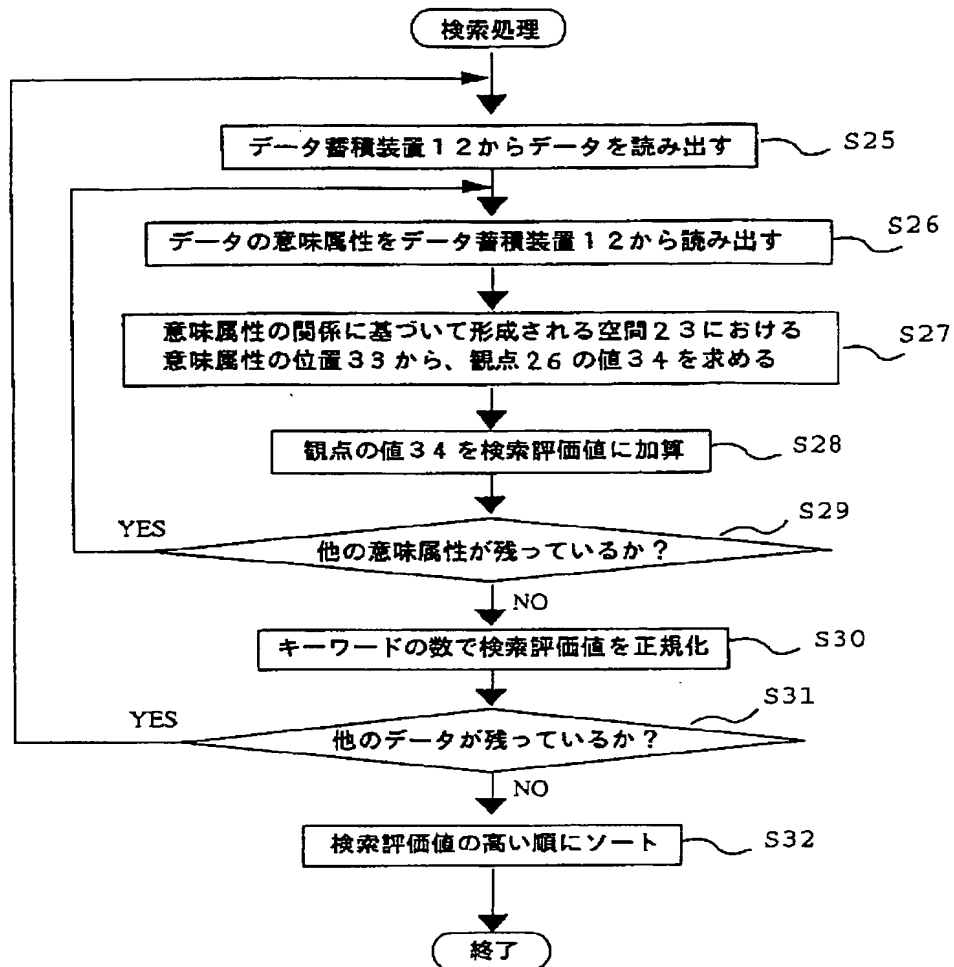
【図 9】



【図 11】



【図 10】



フロントページの続き

(72)発明者 上森 明
東京都千代田区内幸町 1 丁目 1 番 6 号 日
本電信電話株式会社内